Schaefer, Thilo; Peichl, Andreas

Working Paper

Documentation FiFoSiM: integrated tax benefit microsimulation and CGE model

Finanzwissenschaftliche Diskussionsbeiträge / Finanzwissenschaftliches Forschungsinstitut an der Universität zu Köln, No. 06-10

Provided in cooperation with:
Universität zu Köln

FiFo-CPE Discussion Papers
Finanzwissenschaftliche Diskussionsbeiträge

Discussion Paper No. 06-10

Documentation FiFoSiM: Integrated tax benefit microsimulation and CGE model

Andreas Peichl and Thilo Schaefer

2006

Finanzwissenschaftliches Forschungsinstitut an der Universität zu Köln

FiFo Köln is a Member of CPE -
Cologne Center for Public Economics

fifo-koeln.de
cpe.uni-koeln.de

FiFo-CPE Discussion Papers
Finanzwissenschaftliche Diskussionsbeiträge

ISSN 0945-490X

No. 06-10 / October 2006

Documentation FiFoSiM: Integrated tax benefit microsimulation and CGE model*

Andreas Peichl[1]

Thilo Schaefer[2]

* The authors would like to thank Christian Bergs, Erika Berthold, Stephan Dobroschke, Marios Doulis, Clemens Fuest and Sven Heilmann for their helpful contributions. The usual disclaimer applies.

[1] CPE - Cologne Center for Public Economics, University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany. E-Mail: a.peichl@uni-koeln.de.

[2] CPE - Cologne Center for Public Economics, FiFo at the University of Cologne, Zülpicher Str. 182, 50937 Cologne, E-Mail: schaefer@fifo-koeln.de.

Finanzwissenschaftliches Forschungsinstitut
an der Universität zu Köln

A Member of CPE -
Cologne Center for Public Economics

P.O. Box 420520, D-50899 KÖLN
Zülpicher Str. 182, D-50937 KÖLN

T. +49-221-426979
F. +49-221-422352

http://fifo-koeln.de
http://cpe.uni-koeln.de

Documentation FiFoSiM: Integrated tax benefit microsimulation and CGE model

by

Andreas Peichl* and Thilo Schaefer§

October 2006

Abstract

This documentation describes FiFoSiM, the integrated tax benefit microsimulation and CGE model of the Center of Public Economics at the University of Cologne.

FiFoSiM consists of three main parts. The first part is a static tax benefit microsimulation module. The second part adds a behavioural component to the model: an econometricaly estimated labour supply model. The third module is a CGE model which allows the user of FiFoSiM to assess the global economic effects of policy measures.

Two specific features distinguish FiFoSiM from other tax benefit models. First, the simultaneous use of two databases for the tax benefit module and second, the linkage of the tax benefit model to a CGE model.

JEL Codes: D58, H2, J22

Keywords: FiFoSiM, microsimulation, CGE

Acknowledgement: The authors would like to thank Christian Bergs, Erika Berthold Stephan Dobroschke, Marios Doulis, Clemens Fuest and Sven Heilmann for their helpful contributions. The usual disclaimer applies.

*CPE - Cologne Center for Public Economics, University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany. E-Mail: a.peichl@uni-koeln.de

§CPE - Cologne Center for Public Economics, University of Cologne, Zuelpicher Str. 182, 50937 Cologne, E-Mail: schaefer@fifo-koeln.de

Contents

List of Tables

List of Figures

1 Introduction

This technical working paper describes FiFoSiM[1], the integrated tax benefit microsimulation and CGE model of the Center for Public Economics at the University of Cologne (Finanzwissenschaftliches Forschungsinstitut an der Universität zu Köln (FiFo)[2]). The development of FiFoSiM started in September 2004. The first working version was completed one year later. Since then, the model is steadily being improved during the course of writing new publications based on FiFoSiM.[3] As soon as new data becomes available, it is incorporated in the model's database.

FiFoSiM consists of three main parts. The first part is a static tax benefit microsimulation module. The second part adds a behavioural component to the model: an econometricaly estimated labour supply model. The third module is a CGE model which allows the user of FiFoSiM to assess the global economic effects of policy measures.

Two specific features distinguish FiFoSiM from other tax benefit models. First, the simultaneous use of two databases for the tax benefit module and second, the linkage of the tax benefit model with a CGE model.

The first module of FiFoSiM is a static microsimulation model for the German tax and benefit system using income tax and household survey micro data. The approach of FiFoSiM is innovative insofar as it creates a dual database using two micro data sets for Germany: FAST98 and GSOEP.[4] FAST98 is the income tax micro data scientific use-file 1998 containing

[1] This English documentation is a short version of the detailed German description of FiFoSiM which can be found in Fuest et al. (2005b).

[2] The Research Institute for Public Finance at the University of Cologne (FiFo = Finanzwissenschaftliches Forschungsinstitut an der Universität zu Köln) is a non-profit research body pursuing independent economic research and policy consultancy. FiFo's day-to-day work chiefly comprises autonomously financed, long-term research programmes. These programmes supply the theoretical framework for a range of medium- and short-term, market-financed research projects and consultancy mandates.

Over the last fifty years, FiFo's main research topics have naturally changed in line with the developments in public sector economics and changing political objectives. Nevertheless, some aspects of public finance are always on the agenda, and over the past twenty years the following issues have crystallised into the Institute's long-term research topics: Fiscal theory and policy, theory and instruments of national and international environmental policy, direct and indirect taxation, intergovernmental fiscal relations on regional, national and international level, theory and evaluation of public spending programmes and state aids, regional planning and sustainable regional development, innovation theory and technology policy, municipal finances and privatisation.

FiFo regularly performs short- and medium-term studies in these core areas of expertise. Additional subjects are tackled if they offer a deeper insight or a new perspective on one or more of our 'traditional' research topics.

Though legally not part of the University of Cologne, FiFo is attached to it in a relation of institutionalised co-operation and mutual assistance. For instance, professors of public sector economics at the University are simultaneously directors of FiFo. Further information about FiFo can be found at the institute's website: www.fifo-koeln.de.

[3] See chapter 6 for an overview of several applications of FiFoSiM.

[4] In the last years several tax benefit microsimulation models for Germany have been developed (see for example Peichl (2005) or Wagenhals (2004)). Most of these models use either GSOEP or FAST data. FiFoSiM is so far the first model to combine these two databases.

a 10%-sample of the German federal income tax statistics.[5] FAST98 includes the relevant data from income tax files of nearly 3 million households in Germany. Our second data source, the German Socio-Economic Panel (GSOEP), is a representative panel study of private households in Germany.[6] In 2003 GSOEP consists of more than 12,000 households with more than 30,000 individuals. The simultaneous use of both databases allows for the imputation of missing values or variables in the other dataset.[7]

Figure 1 shows the Basic setup of FiFoSiM .

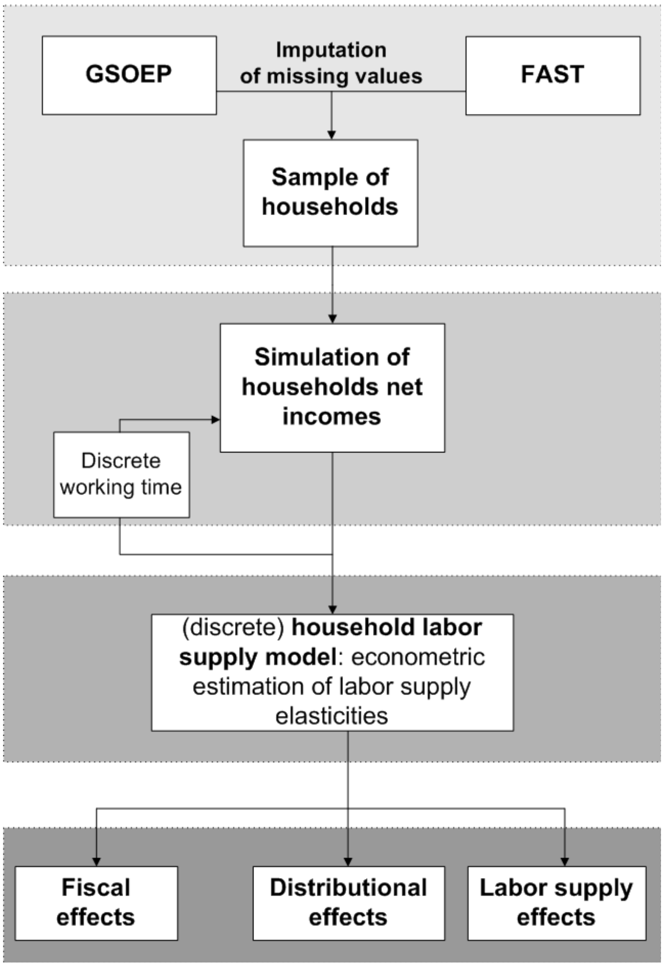

Figure 1: Basic setup FiFoSiM

The layout of FiFoSiM follows several steps: First, the database is updated using the static ageing technique[8] which allows controlling for changes in global structural variables and

[5] Cf. Merz et al. (2005) for a description of FAST98.

[6] Cf. Haisken De-New and Frick (2003) for an introduction to GSOEP.

[7] See Rässler (2002) for an introduction to statistical matching procedures and imputation techniques.

[8] Cf. Gupta and Kapur (2000) for an overview of the techniques to modify the data for the use in microsimulation models.

a differentiated adjustment for different income components of the households. Second, we simulate the current tax system in 2006 using the modified data. The result of this simulation is the benchmark for different reform scenarios which are also modelled using the modified database.

The modelling of the tax and transfer system uses the technique of microsimulation.[9] FiFoSiM computes individual tax payments for each case in the sample considering gross incomes and deductions in detail. The individual results are multiplied by the individual sample weights to extrapolate the fiscal effects of the reform with respect to the whole population. After simulating the tax payments and the received benefits we can compute the disposable income for each household. Based on these household net incomes we estimate the distributional and the labour supply effects of the analysed tax reforms. For the econometric estimation of labour supply elasticities, we apply a discrete choice household labour supply model. Furthermore, FiFoSiM contains a CGE module for the estimation of growth and employment effects, which is linked to the tax benefit module. This interaction allows for a better calibration of the model parameters and a more accurate estimation of the various effects of reform proposals.

The setup of this documentation is as follows. Chapter 2 describes (the creation of) the dual database of FiFoSiM, while chapter 3 describes the tax benefit module. Chapter 4 contains a description of the labour supply model, while chapter 5 describes the CGE module. In chapter 6 several applications of FiFoSiM are presented. Chapter 7 concludes and gives an outlook to some developments planned for the further improvement of FiFoSiM.

[9] Cf. Gupta and Kapur (2000) or Harding (1996) for an introduction to the field of microsimulation.

2 Database

The approach of FiFoSiM is innovative insofar as it creates a dual database using two micro datasets for Germany. The first one, FAST98, consists of micro data from the German federal income tax statistics. Our second data source, the German Socio-Economic Panel (GSOEP), is a representative panel study of private households in Germany. A specific feature of FiFoSiM is the simultaneous use of both databases allowing for the imputation of missing values or variables in the other dataset. Due to the time lags between the census and the availability, the data has to be updated to represent the German economy in the period of analysis. The data sources, the matching and the ageing are described in detail in the following.

2.1 Income tax scientific use-file 1998 (FAST98)

FAST98 is the income tax scientific use-file 1998 (FAST98) containing a 10%-sample of the German federal income tax statistics.[10] FAST98 includes the relevant data from income tax files of nearly 3 million households in Germany.

The federal income tax statistics is published every three years but with a time lag of five to six years. This statistics contains all information from the personal income tax form (e.g. source and amounts of incomes, deductions, age, children) for every household subject to income taxation in Germany. For 1998, almost 30 million households are included in the data base.

The FAST micro data is especially suitable for a detailed analysis of the German tax system. All structural characteristics of the taxpayers are well represented and can be modelled for a differentiating analysis of tax reforms.

2.2 German Socio-Economic Panel (GSOEP)

The German Socio-Economic Panel (GSOEP) is a representative panel study of private households in Germany since 1984.[11] In 2003 GSOEP consists of more than 12,000 households with more than 30,000 individuals. The data include information on earnings, employment, occupational and family biographies, health, personal satisfaction, household composition and living situation.

The panel structure of GSOEP allows for longitudinal and cross section analysis of economic and social changes. Bork (2000) certifies GSOEP a rather good mapping of labour income whereas capital and business income are not represented just as well.

GSOEP contains information about the working time and the social environment of the households which is used for the labour supply estimations. Furthermore, the bottom end of

[10] Cf. Merz et al. (2005) for a description of FAST98.

[11] See SOEP Group (2001) or Haisken De-New and Frick (2003) for a more detailled introduction to GSOEP.

the income distribution is better represented in GSOEP than in FAST.

2.3 Creating the dual database

One special feature of FiFoSiM is the creation and usage of a dual database. To be more precise, FiFoSiM actually consists of two tax benefit microsimulation models. The first one is based on administrative tax data (FAST), the second on household survey data (GSOEP). The main reason for using the dual database instead of having only one merged database is the huge difference in the number of observations (3 million vs. 30,000). Furthermore, both databases have several shortcomings, as described in the previous sections, but nevertheless, they are the two most appropriate datasets available for the analysis of the German tax benefit system. Therefore, information from one database is used for the imputation of missing values or variables in the second dataset and vice versa. A complete matching of the two databases is also possible but not yet necessary as we only need some of the variables from the second file, which are missing for our analysis in the first file.[12] Hence, the dual database of FiFoSiM actually consists of two enhanced datasets, which allow for a better analysis of tax benefit reforms than the two raw datasets. Another aspect is the handling of missing values in existing variables in each dataset. There exist several principal ways for matching datasets or the imputation of missing values.[13] Those used in FiFoSiM are described in the following together with information about the respective implementation.

2.3.1 Imputation of missing values

For the imputation of missing values in one variable several concepts exist.[14] In general, the imputation of missing values stands for replacing missing data with "plausible values"[15]. Let K be a variable from a dataset A with i non-missing values $N = (n_1, n_2, ..., n_i)$ and j missing values $M = (m_1, m_2, ..., m_j)$: $K = (N, M) = (n_1, n_2, ..., n_i, m_1, m_2, ..., m_j)$, and $O = (O_1, O_2, ...)$ a vector of (other) variables without missing values, and H be the same variable as K and P the same as O but from a different dataset B.

[12] There are mainly legal privacy issues in Germany militating against a complete match. Nevertheless, the matching of the anonymised databases does not allow for a deanonymisation of the individuals in the datasets.

[13] This section is based on Rässler (2002), who gives an introduction to statistical matching procedures and imputation techniques, as well as an overview of the literature and software packages.

[14] Cf. Rubin (1987) or Little and Rubin (1987) as additional references for the imputation of missing values. The best but of course most expensive way to impute missing values would be to collect further information on the missing data. But even this solution, however, cannot compensate for shortcomings in historic datasets.

[15] Schafer (1997), p. 1. The alternative to this imputation approach would be to delete (or at least omit) the cases containing missing values. This procedure would lead to biased estimations if the people with missing values share the same characteristics.

Mean substitution In this approach, the missing values M in variable K are either substituted by the mean of the non missing values N:

$$\widehat{K} = (N, \overline{N}) = (n_1, n_2, ..., n_i, \overline{n}, \overline{n}, ..., \overline{n}),$$

or they are substituted by the mean of a similar variable H from a different dataset B:

$$\widehat{K} = (N, \overline{H}) = \left(n_1, n_2, ..., n_i, \overline{h}, \overline{h}, ..., \overline{h}\right),$$

If the missing values can be attributed to some specific subgroups, then the missing values for each subgroup are replaced by the mean of each subgroup either from the non missing values or a different dataset.

This procedure reduces the variance of this variable and should therefore be the last option and only considered if other approaches are not applicable. The latter could be the case if there is, for example, no correlation between the variable containing missing values and any other variable. This approach is no longer used in FiFoSiM.

Regression In the regression approach, a function for the estimation of the missing values is constructed. A (linear) regression[16] of the other (non missing) variables O on the non missing values of K, N, is done:

$$N = O\beta.$$

Or, as in the case of mean substitution, the similar variable H from a different dataset B is regressed on the other variables P from B:

$$H = P\beta.$$

Often a stochastic random value $\widehat{u}$ is added to the prediction of the missing values M to allow for more variation:

$$\begin{aligned} \widehat{M} &= O\widehat{\beta} + \widehat{u}, \\ &or \\ \widehat{M} &= P\widehat{\beta} + \widehat{u}. \end{aligned}$$

These estimates $\widehat{M}$ are then used to replace the missing values N:

[16] For categorical variables often logistic regressions are undertaken. A good textbook introduction to the different regression techniques can be found in Greene (2003).

$$K = \left(N, \widehat{M}\right)$$

In FiFoSiM this approach is mainly used for variables originally coming from the FAST-Database. Most of these missing values are due to anonymisation and their values can be restricted to some intervals due to different information.

Multiple imputation In the multiple imputation approach, multiple values for each missing value are simulated (and hence multiple datasets are generated) to better reflect the variation in the estimates and the uncertainty in the imputation procedure itself:

$$\widetilde{M}^i = (\widetilde{m}_1^i, \widetilde{m}_2^i, ..., \widetilde{m}_j^i)$$

Then the average of these estimates for each observation is calculated as the estimator for the missing values:

$$\widehat{M} = \frac{1}{i} \sum_i \widetilde{M}^i$$

and is used to replace the missing value in the original dataset:

$$K = \left(N, \widehat{M}\right) = \left(n_1, n_2, ..., n_i, \widehat{m}_1^i, \widehat{m}_2^i, ..., \widehat{m}_j^i\right)$$

This approach is used in FiFoSiM for most of the GSOEP variables containing missing values. The relatively small number of cases in the GSOEP allows the use of several simulation runs for the imputation in a few minutes, whereas for the FAST data this method takes noticeably longer.

2.3.2 Statistical matching

The idea of combining two existing datasets to create a joint dataset was developed during the 1970s.[17] The general principle is to merge two (or more) separate databases through the matching of the individual cases. This matching is done on common variables that exist in both databases (for example gender, age and income). Figure 2 illustrates this basic idea of statistical matching.

To put it more analytical[18]: We have three sets of variables X, Y, Z and two samples $A = (X, Y)$ and $B = (X, Z)$. X are the common variables in both samples, Y and Z are sample specific. We can now create a new, joint sample $C = (X, Y, Z)$ by merging a recipient sample

[17] Cf. Okner (1972), Radner et al. (1980) or Cohen (1991).
[18] This is based on Sutherland et al. (2002).

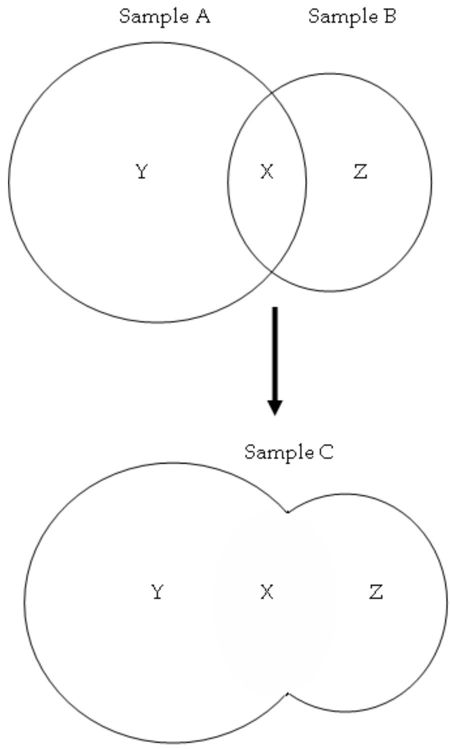

Figure 2: Basic idea of statistical matching

(lets say A) with observations from a donor sample (B) with exact (or close) values of X.[19] In doing so, one assumes the Conditional Independence Assumption (CIA)[20] holds: Conditionally on X, Y and Z are independent.[21]

Of course, one would like to find perfect matches all of the time.[22] But without corresponding identification numbers and large numbers of variables, a perfect match may not always be possible.[23] In our case, an exact matching is not possible, therefore we have to use methods of statistical matching to match close (instead of exact) observations that share a set of common

[19] Which sample should be taken as the recipient, and which as the donor, depends on the particular matching question.

[20] See Sims (1972a), Sims (1972b) and Sims (1974). The CIA means that the X variables contain all information about the relationship between Y and Z. If we know X, Y (Z) contains no additional information about Z (Y).

[21] This can "in practice [...] rarely be checked" (Sutherland et al. (2002)). If the CIA does not hold, one can still use methods of statistical matching if the relationship between Y and Z can be estimated from other sources and incorporated into the matching process (see Paass (1986)).

[22] This would be possible, if one would have variables (name, address, date of birth, social security number) which uniquely identify an individual. Due to privacy reasons it seems impossible for researchers to gain access to raw micro data which include these information without anonymisation. Nevertheless, an exact matching of records from different datasets is possible for those institutions (administration, data collectors), having access to variables which uniquely identify an observation in both datasets.

[23] If many common variables are continuous, a perfect match seems to be impossible (see Rässler (2002), p.18).

characteristics. The idea underlying this matching approach is that if two people have a lot of things in common (like for example age, sex, income, marital status, number of children), then they are likely to have other characteristics (like for example expenses) in common. The statistical matching of two databases can either be done by regression or by methods of data fusion.

Regression In the regression approach, the vector of common variables X is regressed on the specific variables from the donor dataset Z:

$$Z = X\beta.$$

The estimated coefficients β are then used to predict the values of Z in the joint dataset:

$$C(X, Y, \widehat{Z}(\beta)).$$

A strong correlation between X and Z is important for a successful merging. This approach is rather easy to perform, but it has the drawback that information in terms of variation is lost in the second dataset.

Data fusion The data fusion approach can be distinguished into nearest neighbour approach and propensity score matching. The general idea of both approaches is similar, they only differ in the first step.

The first step in the *nearest neighbour* approach is to weight and norm the common variables, whereas in the *propensity score* approach[24], the propensity score is estimated. To do so, a dummy variable I is introduced into the pooled dataset D, containing the common variables X from both samples A, B, indicating 1 if the observation is from the recipient dataset and 0 if it is from the donor dataset:

$$I = \begin{cases} 1 & \text{if observation is from the recipient file} \\ 0 & \text{if observation is from the donor file} \end{cases}$$

Then a logit or probit estimation of the probability of the observation being from the recipient sample (that is of the dummy indicator variable being 1) conditional on the common variables X is done:

[24] Cf. Rosenbaum and Rubin (1983). In general, the propensity score is defined as the conditional probability of treatment given (the common) background variables. Therefore, the propensity score is used as a predictor of the probability of being in the treatment group versus being in the control group. In our case, an observation is in the treatment (control) group if it origins from the recipient (donor) sample.

$$P(I = 1|X) = f(X\beta).$$

The function $f(X\beta)$ is called the propensity score and indicates the probability of the observation belonging to the treatment group (the recipient sample).

The second step is similar for both approaches. The distance between the observations from both datasets is computed using a distance function[25]. In the nearest neighbour case, the distance is based on the weighted common variables, in the propensity score case, the distance is based on the estimates for the propensity scores, which can be interpreted as some sort of implicit weighting function.

In the third step, the joint database C is created by merging the observations from the two datasets A and B with the minimal distance between them.

In FiFoSiM several of these approaches are used due to the difference in the number of observations (3 million vs. 30,000). In general, information from the smaller GSOEP dataset is matched to the FAST data using the regression approach. FAST information is merged to GSOEP data using propensity score matching. Missing values in both datasets are imputed using different approaches depending on the specific circumstances in each case.

The creation of this dual or enhanced database with information from administrative tax data and a household survey gives the users of FiFoSiM a powerful tool for the analysis of various questions regarding the German tax benefit system.

2.4 Updating the data samples

The database is updated using the static ageing technique[26] which allows controlling for changes in global structural variables and a differentiated adjustment for different income components of the households. Especially the income tax data sample needs to be updated as it describes the situation of 1998. The GSOEP data only needs to be adjusted from 2002.

The first step is to reproduce the fundamental structural changes of the population. This is done according to the following criteria: age (in 5 year categories), assessment for income tax (separate or joint) and region (East/West Germany). The method applied here follows

[25] See Cohen (1991). In general, three different distance functions can be used to determine similarity between the two samples: the absolute, Euclidean or Mahalanobis distance. Let x_i^A denote the common variables of unit i in sample A and x_j^B those of unit j in sample B.

The absolute distance is defined as $d_{ij}^{abs} = \left(x_i^A - x_j^B\right)$.

The Euclidean distance is given by $d_{ij}^{E} = \sqrt{\left(x_i^A - x_j^B\right)' \left(x_i^A - x_j^B\right)}$.

The Mahalanobis distance (see Mahalanobis (1936)) is based on the correlation matrix S_X^{-1} between the two sets of variables: $d_{ij}^{M} = \sqrt{\left(x_i^A - x_j^B\right)' S_X^{-1} \left(x_i^A - x_j^B\right)}$.

[26] Cf. Gupta and Kapur (2000) for an overview of the techniques to modify the data for the use in microsimulation models.

Quinke (2001): The cases from the FAST sample are compared to aggregated statistical data for the whole population regarding the above named criteria to calculate the degree of coverage. Assuming that this degree remains stable over the years, the actual aggregate population statistics and prognosis for the year 2006 times the coverage degree allows for an approximate adjustment of the database to account for the basic structural changes. Technically, the sample weights need to be adjusted. The weighting coefficients indicate how many actual cases of the real population are represented by each case in the sample. Using the software package Adjust by Merz et al. (2001) the sample weights are adjusted according to 52 possible combinations of the attributes (13 age categories times 2 assessment types times 2 regions). Now, the extrapolation of the sample using the adjusted weights represents the actual population structure better.

In the second step, the taxpayer's incomes are updated with respect to the varying development of different income types. Also different income growth rates between West and East as well as decreasing negative incomes are taken into account. This allows for a differentiated estimation of the income development. Based on empirical research of the DIW[27] different coefficients for positive and negative incomes are applied on each case's income. For the simulation model this means that each income value is multiplied with the specific coefficient and thus extrapolated to the current income level. Of course, the coefficients only represent the average development, but regarding the whole population this method provides a satisfying approximation to the income structure of today.

[27] Cf. Bach and Schulz (2003).

3 Tax benefit module

In this section, the modelling of the German tax benefit system is described.[28]

3.1 Modelling the German income tax law 2006

Individuals are subject to personal income tax. Residents are taxed on their global income; non-residents are taxed on income earned in Germany only.

3.1.1 Income sources

The basic steps for the calculation of the personal income tax under German tax law are according to the scheme of table 1 as follows. The first step is to determine a taxpayer's income from different sources and to allocate it to the seven forms of income. The German tax law distinguishes between seven different categories of income: income from agriculture and forestry, business income, self employment income, salaries and wages from employment, investment income, rental income and other income (including, for example, annuities and certain capital gains). For each type of income, the tax law allows for certain income related deductions. In principle, all expenses that are necessary to obtain, maintain or preserve the income from a source are deductible from the receipts of that source. The second step is to sum up these incomes to obtain the adjusted gross income. Third, deductions like contributions to pension plans or charitable donations are taken into account, which gives taxable income as a result. Finally, the income tax is calculated by applying the tax rate schedule to taxable income.

	Sum of net incomes from 7 categories (receipts from each source minus expenses)
=	adjusted gross income
-	deductions (social security and insurance contributions, personal expenses)
=	taxable income x
$\cdot$	tax formula
=	tax payment T

Table 1: Calculation of the personal income tax

[28] As the Germany tax benefit system is very complex, we focus on the major parts of the model in this description. A more detailed description can be found in the German version of this documentation (see Fuest et al. (2005b)).

3.1.2 Taxable income

The subtraction of special expenses (*Sonderausgaben*) and expenses for extraordinary burden (*außergewöhnliche Belastungen*) from adjusted gross income gives taxable income.

The special expenses consist of:

- alimony payments (maximum of 13,805 € per year)
- church tax
- tax consultant fees
- expenses for professional training (up to 4,000 € per year)
- school fees of children (up to 30%)
- charitable donations (up to 5% of the adjusted gross income)
- donations to political parties (up to 1,650 €)
- expenses for financial provision, i.e. insurance premiums (pension schemes up to 20,000 € per person, health/nursing care/unemployment insurance, life assurance, disability insurance)

The insurance contributions are normally equally split between employer and employee. Each premium is calculated as contribution rate times the income that is subject to contributions up to the according contribution ceiling. Current contribution rates are 19.5% for old age insurance (5,200 € ceiling in West Germany / 4,400 € in East Germany), (an assumed average of) 13.25% for health insurance (3,525 € ceiling), 6.5% for unemployment insurance (ceilings: 5,200 €/4,400 €) and 1.7% for nursing care insurance (same ceiling as health insurance) plus various special supplements.

The expenses for extraordinary burden consist of:

- expenses for the education of dependants, expenses for the cure of illness, expenses for home help with elderly or disabled people, commuting expenses caused by disability in certain cases
- allowances for disabled persons, surviving dependants and persons in need of care
- child care costs
- tax allowances for self used proprietary, premises and historical buildings

Loss deduction according to § 10d EStG

1. Deductions of negative income up to 511,500 € income of the preceding assessment period [loss carried back]

2. Deductions of negative income up to 1 million Mio. € and 60% above of income of the following assessment period [loss carried forward]

In the FAST calculation, loss deduction is explicitly taken into account with the extrapolated data, while the GSOEP version has to use the FAST results as there is no according information in the GSOEP data.

Income Sum of income minus the aforementioned deductions gives the income. This may be reduced by the equitable compensation (§ 46 Abs. 3 EStG) which is a variable in the FAST data. The FAST data also accounts for the household allowance which is no longer in force but it indicates if an allowance for single parents can be applied as the latter is the follow-up regulation.

Child allowance cf. §§ 31, 32 Abs. 6 EStG

Child allowance (2904 € per parent deduction from taxable income) or child benefit (154 € per month for the 1st to 3rd child, 179 € as from the 4th child). If the child allowance is more favourable, it is deducted from the taxable income while the sum of child benefits is added to the tax due.

The model includes this regulation as it compares allowance and benefit for each case. The FAST version accounts for 4 children only.

The possible deduction of child allowances gives the **taxable income**.

3.1.3 Tax due

To determine the tax due exemptions with progression (§ 32b EStG) and certain special calculation methods (§§ 34, 34b EStG) have to be applied which only the FAST data accounts for.

The tax is calculated on the basis of a mathematical formula which, as of the year 2004, is structured as follows[29]:

[29] The basic structure of the underlying progressive tax schedule according to the current tax law is: $T(x) =$

$$\begin{cases} 0 & \text{if } x \leq G \\ \left(\frac{t_m - t_e}{2(M-G)}(x-G) + t_e\right)(x-G) & \text{if } G < x \leq M \\ \left(\frac{t_s - t_m}{2(S-M)}(x-M) + t_m\right)(x-M) + (M-G)\frac{t_m + t_e}{2} & \text{if } M < x \leq S \\ t_s(x-S) + \frac{t_s + t_m}{2}(S-M) + \frac{t_m + t_e}{2}(M-G) & \text{if } x > S \end{cases}$$

$$EStt = \begin{cases} 0 & if\ zvE \leq 7664 \\ (883,74 \cdot \frac{zvE-7664}{10000} + 1500) \cdot \frac{zvE-7664}{10000} & if\ 7664 < zvE \leq 12739 \\ (228,74 \cdot \frac{zvE-12739}{10000} + 2397) \cdot \frac{zvE-12739}{10000} + 989 & if\ 12739 < zvE \leq 52151 \\ 0,42 \cdot zvE - 7914 & if\ zvE > 52151 \end{cases}$$

For married taxpayers filing jointly, the tax is twice the amount of applying the formula to half of the married couple's joint taxable income.

3.2 Modelling the benefit system

To simulate the labour supply effects, the calculation of net incomes has to take the transfer system into account as well. State transfers such as unemployment benefit, housing benefit, and social benefits are modelled in FiFoSiM.

3.2.1 Unemployment benefit I

Law Persons seeking work who were employed subject to social insurance contributions at least 12 months before getting unemployed are entitled to receive the so-called unemployment benefit I (according to the German SGB III). The amount to be paid depends on the gross income on a certain date. This is reduced by 21% for social contributions and the individual income tax. The unemployment benefit I amounts to 60% of the resulting net income (or 67% for unemployed with children).

The benefit period depends on age and seniority (as shown in the following table 2).

old regulation until 31.01.2006			**new regulation from 01.02.2006**		
employment	age	benefit period	employment	age	benefit period
12		6	12		6
16		8	16		8
20		10	20		10
24		12	24		12
30	45	14	30	55	15
36	45	18	36	55	18
44	47	22			
52	52	26			
64	57	32			

Table 2: Duration of unemployment benefit entitlement

where, x indicates the tax base, $T(x)$ the tax payment, G is the basic personal allowance, M the upper limit of the first progression zone, S the lower limit applicable to the top rate t_s, t_e the lowest tax rate and t_m the highest tax rate of the lower progression zone (i.e. the lowest tax rate of the upper progression zone).

Model The GSOEP panel data contains information about previous unemployment benefit payments, employment periods, etc. When modelling a person's working time categories it has to be examined whether the person might get unemployment benefits in certain working time categories. This is assumed for persons who received unemployment benefits or who were employed subject to social insurance contributions at least 12 month within the last 36 month. The amount of benefit paid is calculated as described above. The remaining net income is deducted from the unemployment benefit.

3.2.2 Unemployment benefit II

Law The unemployment benefit II replaced the former system of unemployment support and social benefits in the course of the so-called Hartz reform. All employable persons between 15 and 65 years and the persons living with them in the same household are entitled to receive unemployment benefit II, as soon as they are no longer entitled to receive unemployment benefitI.

In contrast to the latter, unemployment benefit II depends on the neediness of the recipient and is therefore means-tested. Needy is a person who, by its own household's income, is not able to satisfy the own elementary needs and those of the persons living in his household. The unemployment benefit II corresponds to the former social benefits plus housing and heating costs if necessary.

Model Unemployment benefit II is modelled according to the former social benefits. The basic amount for each person counts as the need which is means-tested against the household's net income.

3.2.3 Social benefits

Law Persons who are not able to take care of their subsistence are entitled to receive social benefits. Since unemployment benefit II (see above) was introduced, only non employable persons can receive social benefits. Further on, social benefits are paid in extraordinary circumstances such as impairment of health.

Model Analogously to unemployment benefit II the basic amount for each person and their respective household net income are taken into account to determine the amount of social benefits actually paid. Social benefits for persons in extraordinary circumstances are not reflected in the model due to missing information in the data.

3.2.4 Housing benefits

Law Housing benefits are paid on request to tenants as well as to owners. The number of persons living in the household, the number of family members, the income and the rent depending on the local rent level determine if a person is entitled to receive housing benefits.

Model First, summing up the individual incomes considering the basic allowances gives the chargeable household income. Then, due to missing information about local rent levels, the weighted averages of rents up to the maximum support allowed are taken into account to determine the housing benefits.

4 Labour supply module

To analyse the behavioural responses induced by the different tax reform scenarios we simulate their labour supply effects. Following Van Soest (1995) we apply a discrete choice household labour supply model,[30] assuming that the household's head and his partner jointly maximise a household utility function in the arguments leisure of both partners and net income. Household i $(i = 1, ..., N)$ can choose between a finite number $(j = 1, ..., J)$ of combinations $(y_{ij}, lm_{ij}, lf_{ij})$, where y_{ij} is the net income, lm_{ij} the leisure of the husband and lf_{ij} the leisure of the wife of household i in combination j. Based on our data we choose three working time categories for men (unemployed, employed, overtime) and five for women (unemployed, employed, overtime and two part time categories).

We model the following translog[31] household utility function

$$V_{ij}(x_{ij}) = x_{ij}' A x_{ij} + \beta' x_{ij} \tag{1}$$

where $x = \begin{pmatrix} \ln y_{ij}, & \ln lm_{ij}, & \ln lf_{ij} \end{pmatrix}'$ is the vector of the natural logs of the arguments of the utility function. The elements of x enter the utility function in linear (coefficients $\beta = (\beta_1, \beta_2, \beta_3)'$) and in quadratic and gross terms (coefficients $A_{(3\times3)} = (a_{ij})$). Using control variables z_p $(p = 1, ..., P)$[32] we control for observed heterogeneity in household preferences by defining the parameters β_m, α_{mn} as

$$\beta_m = \sum\nolimits_{p=1}^{P} \beta_{mp} z_p \tag{2}$$

$$\alpha_{mn} = \sum\nolimits_{p=1}^{P} \alpha_{mnp} z_p \tag{3}$$

where $m, n = 1, 2, 3$.

Following McFadden (1973) and his concept of random utility maximisation[33] we add a

[30] A detailed description of the FiFoSiM labour supply module can be found in Fuest et al. (2005b). A survey of different kinds of labour supply models is provided by Blundell and MaCurdy (1999), Creedy et al. (2002) and Hausman (1985) especially for continuous models. Using a discrete choice model has the advantage of the possibility to model nonlinear budget constraints (see Van Soest (1995) or MaCurdy et al. (1990)). Furthermore a discrete choice between distinct categories of working time seems to be more realistic as a continuum of choices because of working time regulations.

[31] Cf. Christensen et al. (1971).

[32] We use control variables for age, children, region and nationality , which are interacted with the leisure terms in the utility function because variables without variation across alternatives drop out of the estimation in the conditional logit model (see Train (2003)).

[33] Cf. McFadden (1981), McFadden (1985) and Greene (2003).

stochastic error term ε_{ij} for unobserved factors to the household utility function:

$$\begin{aligned} U_{ij}(x_{ij}) &= V_{ij}(x_{ij}) + \varepsilon_{ij} \\ &= x'_{ij}Ax_{ij} + \beta' x_{ij} + \varepsilon_{ij} \end{aligned} \tag{4}$$

Assuming joint maximisation of the households utility function implies that household i chooses category k if the utility index of category k exceeds the utility index of any other category $l \in \{1, ..., J\} \backslash \{k\}$, if $U_{ik} > U_{il}$. This discrete choice modelling of the labour supply decision uses the probability of i to choose k relative to any other alternative l:

$$P(U_{ik} > U_{il}) = P[(x'_{ik}Ax_{ik} + \beta' x_{ik}) - (x'_{il}Ax_{il} + \beta' x_{il}) > \varepsilon_{il} - \varepsilon_{ik}] \tag{5}$$

Assuming that ε_{ij} are independently and identically distributed across all categories j to a Gumbel (extreme value) distribution, the difference of the utility index between any two categories follows a logistic distribution. This distributional assumption implies that the probability of choosing alternative $k \in \{1, ..., J\}$ for household i can be described by a conditional logit model[34]:

$$\begin{aligned} P(U_{ik} > U_{il}) &= \frac{\exp(V_{ik})}{\sum_{l=1}^{J} \exp(V_{il})} \\ &= \frac{\exp(x'_{ik}Ax_{ik} + \beta' x_{ik})}{\sum_{l=1}^{J} \exp(x'_{il}Ax_{il} + \beta' x_{il})} \end{aligned} \tag{6}$$

For the maximum likelihood estimation of the coefficients we assume that the hourly wage is constant across the working hour categories and does not depend on the actual working time.[35] For unemployed people we estimate their (possible) hourly wages by using the Heckman correction for sample selection[36]. The household's net incomes for each working time category are computed in the tax benefit module of FiFoSiM.

The labour supply module of FiFoSiM is based on GSOEP data, which is enriched by information taken from the FAST data as described in section 2.3. The sample of tax units is then categorised into 6 groups according to their assumed labour supply behaviour. We distinguish fully flexible couple households (both spouses are flexible), two types of partially flexible couple households (only the male or the female spouse has a flexible labour supply), flexible female and flexible male single households, and inflexible households. We assume that

[34] McFadden (1973). Cf. Greene (2003) or Train (2003) for textbook presentations.

[35] Cf. Van Soest and Das (2001).

[36] Cf. Heckman (1976) and Heckman (1979). A detailed description of these estimations can be found in Fuest et al. (2005b).

a person is not flexible in its labour supply, meaning he or she has an inelastic labour supply, if a person is either

- younger then 16 or older then 65 years of age,
- in education or military service
- receiving old-age or disability pensions
- self employed or civil servant.

Every other employed or unemployed person is assumed to have an elastic labour supply. We distinguish between flexible and inflexible persons, because the labour supply decision of those assumed to be inflexible (e.g. pensioners, students) is supposed to be based on a different consumption leisure decision (or at least with a different weighting of the relevant determinants[37]) than that of those working full time.

[37] Therefore, it is not possible to assume the same econometric relationship for these persons.

5 CGE module

The computable general equilibrium module of FiFoSiM allows us to simulate the overall economic effects of policy changes.[38] The static CGE module of FiFoSiM models a small open economy with 12 sectors and one representative household. The CGE module is programmed in GAMS/MPSGE (Rutherford (1999), Brooke et al. (1998)).

5.1 The model

5.1.1 Households

The representative household maximises a nested CES utility function according to figure 3.

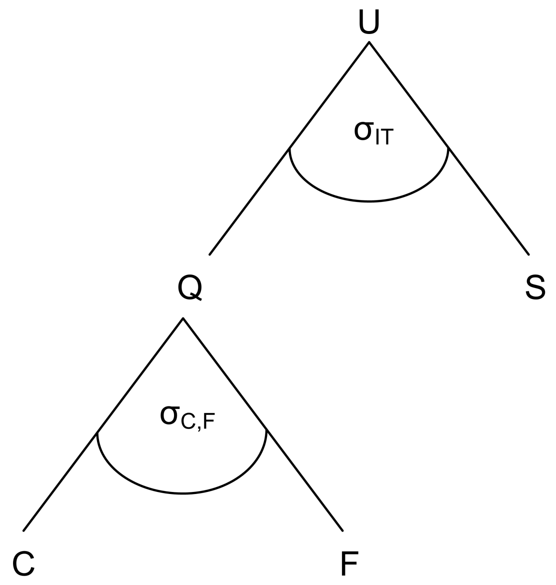

Figure 3: Household level FiFoSiM

At the top nest it maximises its intertemporal utility function $U\ (Q,S)$ choosing between aggregated consumption (including leisure) today Q or in the future S. The result of this optimisation is the savings supply. On the second level, the present consumption leisure (or labour leisure) decision takes place. The household maximises a CES utility function:

$$U\ (C,F) = \left[(1-\beta)^{\frac{1}{\sigma_{C,F}}} C^{\rho_{C,F}} + \beta^{\frac{1}{\sigma_{C,F}}} F^{\rho_{C,F}}\right]^{\frac{1}{\rho_{C,F}}}. \tag{7}$$

where β is the value share, and $\sigma_{C,F} = \frac{\rho_{C,F}-1}{\rho_{C,F}}$ the elasticity of substitution between consumption

[38] This section is based on Bergs and Peichl (2006).

and leisure. The budget constraint is:

$$p^C C = w\left(1 - t^l\right)(E - F) + r\left(1 - t^k\right)K + \overline{T}_{LS}. \tag{8}$$

Consumption $p^C C$ is financed by labour income $w\left(1 - t^l\right)(E - F)$, capital income $r\left(1 - t^k\right)K$ and the lump sum transfer $\overline{T}_{LS}$, that ensures revenue neutrality. Optimising (7) subject to (8) yields the demand functions for goods and leisure. From the latter we calculate the labour supply of the household.

5.1.2 Firms

A representative firm produces a homogenous output in each production sector according to a nested CES production function. Figure 4 provides an overview of the nesting structure.

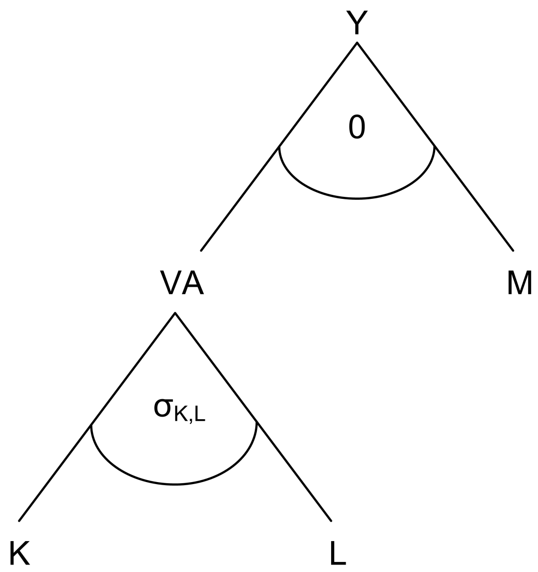

Figure 4: Production structure of FiFoSiM

At the top level nest, aggregate value added (VA) is combined in fixed proportions (Leontief production function) with a material composite (M). M consists of intermediate inputs with fixed coefficients, whereas VA consists of labour (L) and capital (K). The optimisation problem at the top level in each sector i can be written as:

$$Y_i = \min\left[\frac{1}{a_{0i}} f_i(L_i, K_i); \frac{M_{1i}}{a_{1i}}; ...; \frac{M_{12i}}{a_{12i}}\right] \tag{9}$$

In the bottom nest, the following CES function is used:

$$f_i(L_i, K_i) = [\alpha_i L_i^{\rho_i} + (1 - \alpha_i) K_i^{\rho_i}]^{\frac{1}{\rho_i}} \tag{10}$$

where $\sigma_i = \frac{1}{1-\rho_i}$ is the constant elasticity of substitution between labour and capital.

The flexible structure of the model allows for different levels of aggregation ranging from 12 to 7 to 3 to 1 sectors.

5.1.3 Labour market

To account for imperfections of the German labour market, a minimum wage $w^{\min}$ is introduced. The labour supply is therefore rationed:

$$L^S (1 - \mu) \geq L^D. \tag{11}$$

The minimum wage is calibrated so that the benchmark represents the current unemployment level of Germany.

5.1.4 Government

The government provides public goods (G), which are financed by input taxes on labour and capital t^l and t^k. A lump sum transfer to the households completes the budget equation:

$$G + \overline{T}_{LS} = t^l wL + t^k rK. \tag{12}$$

5.1.5 Foreign trade

Domestically produced goods are transformed through a CET-function into specific goods for the domestic and the export market, respectively. By the small-open-economy assumption, export and import prices in foreign currency are not affected by the behaviour of the domestic economy. Analogously to the export side, we adopt the Armington assumption[39] of product heterogeneity for the import side. A CES function characterises the choice between imported and domestically produced varieties of the same good. The Armington good enters intermediate and final demand.

5.2 Data and calibration

The model is based on a social accounting matrix for Germany which is created using the 2000 Input-Output-Table[40] and the static ageing technique to transform the data to 2006.

[39] Vgl. Armington (1969).
[40] Vgl. Statistisches Bundesamt (2005).

The elasticities for the utility and production functions are calibrated based on empirical estimations. The sectoral Armington elasticies are based on Welsch (2001), the elasticity of substitution between labour and capital is assumed to be 0.39 according to Chirinko et al. (2004). The elasticity of intertemporal substitution is assumed to be 0.8 (Schmidt and Straubhaar (1996)) as well as the elasticity of substitution between consumption and leisure (Auerbach and Kotlikoff (1987)).

5.3 Linking the microsimulation and the CGE module

During the last years, the trend of linking micro and macro models emerged[41]. The combination of these two model types allows the utilisation of the advantages of both types of models.

There are two general possibilities for linking the models. On the one hand, one can completely integrate both models[42] or on the other hand, one could combine two separated models via interfaces[43]. The first approach requires the complete micro model to be included in the CGE model which demands high standards for the database and the construction of the integrated model. This often results in various simplifying assumptions.

The second approach can be differentiated into „top-down“, „bottom-up“ or „top-down bottom-up“ approach[44]. The top-down approach computes the macroeconomic variables (price level, growth rates) in a CGE model as input for the micro model. The bottom-up approach works the other way around and information from the micro model (elasticities, tax rates) is used in the macro model. Both approaches suffer from the drawback that not all feedback is used.

The top-down bottom-up approach combines both methods to a recursive approach. In an iterative process one model is solved, information are sent to the other model, which is solved and gives feedback to the first model. This iterative process continues until the two models converge. Böhringer and Rutherford (2006) describe an algorithm for the sequential calibration of a CGE model to use the top-down bottom-up approach with micro models with large numbers of households.

FiFoSiM so far uses either the top-down or the bottom-up approach to combine the microsimulation and the CGE module.

[41] Cf. Davies (2004) for an overview. Most of these models deal with trade liberalization in developing countries.

[42] Cf. ie. Cogneau and Robilliard (2000) or Cororaton et al. (2005).

[43] Cf. Bourguignon et al. (2003).

[44] Cf. Savard (2003) or Böhringer and Rutherford (2006).

6 Applications of FiFoSiM

The development of FiFoSiM started in September 2004. The first running version of the whole system was ready for use one yaer later. Since then, the model has been steadily improved and used for writing new publications.

During the development of FiFoSiM, some introductory papers have been written. Peichl (2005) gives an overview on the evaluation of tax reforms using simulation models. Bergs and Peichl (2006) survey the basic principles and possible applications of CGE models. Ochmann and Peichl (2006) give an introduction to the measurement of distributional effects of fiscal reforms.

Furthermore, FiFoSiM can be used in many ways for the analysis of (reforms of) the tax benefit system. Fuest et al. (2005a) analyse the fiscal, employment and growth effects of the reform proposal by Mitschke (2004). In Fuest, Heilmann, Peichl, Schaefer and Bergs (2006) this analysis is expanded to the negative income tax part (*Bürgergeld*).

Fuest, Peichl and Schaefer (2006c) and Fuest, Peichl and Schaefer (2006b) analyse the efficiency and equity effects of tax simplification. Tax simplification is modelled as the abolition of a set of deductions from the tax base included in the German income tax system. Furthermore, Peichl et al. (2006) analyse the effects of these simplification measures on poverty and richness in Germany.

Fuest, Peichl and Schaefer (2006a) analyse the distributional effects of different flat tax reform scenarios for Germany. Bergs et al. (2006b) and Bergs et al. (2006a) analyse different reform proposals for the taxation of families in Germany.

7 Further Development and conclusion

The aim of this paper was to describe FiFoSiM and its features. FiFoSiM consists of three main parts: a static tax benefit micro simulation model, an econometric estimated labour supply model and a CGE model. Two specific features distinguish FiFoSiM from other tax benefit models. First, the simultaneous use of two databases for the tax benefit module and second, the linkage of the tax benefit model with a CGE model. FiFoSiM can be used to analyse various policy reforms of the complex German tax and transfer system.

Nevertheless, several ideas for the further improvement of FiFoSiM exist. One major aspect of improvement is the modelling of indirect taxes. For this reason, expenditure data is needed and a third data source has to be included into the FiFoSiM database. The micro macro linkage between the microsimulation and the CGE module shall be improved using the top down bottom up approach. Furthermore, the CGE module is to be improved as well, for example by allowing for more different household types or a more sophisticated modelling of the labour market. Moreover, dynamic modules are planned. A small Ramsey type dynamic version of the CGE module exists, but has not been used for any publication yet. This module shall be improved and used in the future.

To conclude, FiFoSiM is a state of the art tax benefit model for Germany. The development is not settled yet. We expect new issues of the FAST and GSOEP data, which have to be implemented in the model, soon. Therefore, this documentation will be updated whenever necessary.

References

Armington, P. (1969). A Theory of Demand for Products distinguished by Place of Production, *IMF Staff Papers* **16**: 159–176.

Auerbach, A. and Kotlikoff, L. (1987). *Dynamic Fiscal Policy*, Cambridge University Press, Cambridge.

Bach, S. and Schulz, E. (2003). Fortschreibungs- und Hochrechnungsrahmen für ein Einkommensteuer- Simulationsmodell. Projektbericht 1 zur Forschungskooperation Mikrosimulation mit dem Bundesministerium der Finanzen, Materialien des DIW Berlin, Nr. 26.

Bergs, C., Fuest, C., Peichl, A. and Schaefer, T. (2006a). Das Familienrealsplitting als Reformoption in der Familienbesteuerung, *Wirtschaftsdienst (forthcoming)* .

Bergs, C., Fuest, C., Peichl, A. and Schaefer, T. (2006b). Reformoptionen der Familienbesteuerung - Aufkommens-, Verteilungs- und Arbeitsangebotseffekte, *Jahrbuch für Wirtschaftswissenschaften (Review of Economics) (forthcoming)* .

Bergs, C. and Peichl, A. (2006). Numerische Gleichgewichtsmodelle - Grundlagen und Anwendungsgebiete, Finanzwissenschaftliche Diskusssionsbeiträge Nr. 06-2.

Blundell, R. and MaCurdy, T. (1999). Labor Supply: A Review of Alternative Approaches, *in* O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics, Vol. 3A*, Elsevier, pp. 1559–1695.

Böhringer, C. and Rutherford, T. (2006). Combining Top-Down and Bottom-up in Energy Policy Analysis: A Decomposition Approach, ZEW Discussion Paper No. 06-007.

Bork, C. (2000). *Steuern, Transfers und private Haushalte. Eine mikroanalytische Simulationsstudie der Aufkommens- und Verteilungswirkungen*, Peter Lang, Frankfurt am Main.

Bourguignon, F., Robilliard, A.-S. and Robinson, S. (2003). Representative versus Real Households in the Macro-Economic Modelling of Inequality, DIAL Document de Travail DT/2003-10.

Brooke, A., Kendrick, D., Meeraus, A. and Raman, R. (1998). GAMS - A Users Guide.

Chirinko, R. S., Fazzari, S. M. and Meyer, A. P. (2004). That Elusive Elasticity: A Long-Panel Approach to Estimating the Capital-Labor Substitution Elasticity, CESifo-Working Paper No. 1240.

Christensen, L., Jorgenson, D. and Lau, L. (1971). Conjugate Duality and the Transcedental Logarithmic Function, *Econometrica* **39**: 255–256.

Cogneau, D. and Robilliard, A.-S. (2000). Growth, Distribution and Poverty in Madagascar: Learning from a Microsimulation Model in a General Equilibrium Framework, International Food Policy Research Institute TMD Discussion Paper 61.

Cohen, M. L. (1991). Statistical matching and microsimulation models, *in* C. F. Citro and E. A. Hanushek (eds), *Improving information for social policy decisions: the uses of microsimulation modelling, Vol II Technical Papers*, National Academy Press, Washington D.C., pp. 62–85.

Cororaton, C. B., Cockburn, J. and Corong, E. (2005). Doha Scenarios, Trade Reforms, and Poverty in the Philippines: A Computable General Equilibrium Analysis, World Bank Policy Research Working Paper 3738.

Creedy, J., Duncan, A., Harris, M. and Scutella, R. (2002). *Microsimulation Modelling of Taxation and the Labour Market: the Melbourne Institute Tax and Transfer Simulator*, Edward Elgar Publishing, Cheltenham.

Davies, J. (2004). Microsimulation, CGE and Macro Modelling for Transition and Developing Economies, Mimeo, University of Western Ontario.

Fuest, C., Heilmann, S., Peichl, A., Schaefer, T. and Bergs, C. (2006). Aufkommens-, Beschäftigungs- und Wachstumswirkungen einer Reform des Steuer- und Transfersystems nach dem Bürgergeld-Vorschlag von Joachim Mitschke, FiFo-Bericht 08-2006.

Fuest, C., Peichl, A. and Schaefer, T. (2005a). Aufkommens-, Beschäftigungs- und Wachstumswirkungen einer Steuerreform nach dem Vorschlag von Mitschke, FiFo-Bericht 05-2005.

Fuest, C., Peichl, A. and Schaefer, T. (2005b). Dokumentation FiFoSiM: Integriertes Steuer-Transfer-Mikrosimulations- und CGE-Modell, Finanzwissenschaftliche Diskussionsbeiträge Nr. 05 - 03.

Fuest, C., Peichl, A. and Schaefer, T. (2006a). Die Flat Tax: Wer gewinnt? Wer verliert? Eine empirische Analyse für Deutschland., *Steuer und Wirtschaft (forthcoming)* .

Fuest, C., Peichl, A. and Schaefer, T. (2006b). Does Tax Simplification yield more Equity and Efficiency? An empirical analysis for Germany, Finanzwissenschaftliche Diskussionsbeiträge Nr. 06 - 05.

Fuest, C., Peichl, A. and Schaefer, T. (2006c). Führt Steuervereinfachung zu einer "gerechteren" Einkommensverteilung? Eine empirische Analyse für Deutschland, *Perspektiven der Wirtschaftspolitik (forthcoming)* .

Greene, W. (2003). *Econometric Analysis*, Prentice Hall, New Jersey.

Gupta, A. and Kapur, V. (2000). *Microsimulation in Government Policy and Forecasting*, North-Holland, Amsterdam.

Haisken De-New, J. and Frick, J. (2003). DTC - Desktop Compendium to The German Socio-Economic Panel Study (GSOEP).

Harding, A. (1996). *Microsimulation and public policy*, North-Holland, Elsevier, Amsterdam.

Hausman, J. (1985). Taxes and Labor Supply, *in* A. Auerbach and M. Feldstein (eds), *Handbook of Public Economics*, North-Holland, Amsterdam, pp. 213–263.

Heckman, J. (1976). The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models, *Annals of Economic and Social Measurement* **5**: 475–492.

Heckman, J. (1979). Sample Selection Bias as a Specification Error, *Econometrica* **47**: 153–161.

Little, R. and Rubin, D. (1987). *Statistical Analysis with Missing Data*, John Wiley & Sons, New York.

MaCurdy, T., Green, D. and Paarsch, H. (1990). Assessing Empirical Approaches for Analyzing Taxes and Labor Supply, *Journal of Human Resources* **25(3)**: 415–490.

Mahalanobis, P. (1936). On the generalised distance in statistics, *Proceedings of the National Institute of Science of India* **12**: 49–55.

McFadden, D. (1973). Conditional Logit Analysis of Qualitative Choice Behaviour, *in* P. Zarembka (ed.), *Frontiers in Econometrics*, New York, pp. 105–142.

McFadden, D. (1981). Econometric Models of Probabilistic Choice, *in* C. Manski and D. McFadden (eds), *Structural Analysis of Discrete Data and Econometric Applications*, The MIT Press, Cambridge, pp. 198–272.

McFadden, D. (1985). Econometric Analysis of Qualitative Response Models, *in* Z. Griliches and M. Intrilligator (eds), *Handbook of Econometrics*, Elsevier, Amsterdam, pp. 1396–1456.

Merz, J., Stolze, H. and Imme, S. (2001). ADJUST FOR WINDOWS - A Program Package to Adjust Microdata by the Minimum Information Loss Principle, FFB-Dokumentation No. 9, Department of Economics and Social Sciences, University of Lüneburg, Lüneburg.

Merz, J., Vorgrimler, D. and Zwick, M. (2005). De facto anonymised microdata file on income tax statistics 1998, FDZ-Arbeitspapier Nr. 5.

Mitschke, J. (2004). *Erneuerung des deutschen Einkommensteuerrechts: Gesetzestextentwurf und Begründung*, Verlag Otto Schmidt, Köln.

Ochmann, R. and Peichl, A. (2006). Measuring Distributional Effects of Fiscal Reforms, Finanzwissenschaftliche Diskussionsbeiträge Nr. 06 - 09.

Okner, B. (1972). Constructing a New Data Base from Existing Microdata Sets: The 1966 Merge File, *Annals of Economic and Social Measurement* pp. 325–342.

Paass, G. (1986). Statistical match: Evaluation of existing procedures and improvements by using additional information, *in* G. H. Orcutt and H. Quinke (eds), *Microanalytic Simulation Models to Support Social and Financial Policy*, Elsevier Science, Amsterdam, pp. 401–422.

Peichl, A. (2005). Die Evaluation von Steuerreformen durch Simulationsmodelle, Finanzwissenschaftliche Diskussionsbeiträge Nr. 05-01.

Peichl, A., Schaefer, T. and Scheicher, C. (2006). Poverty and Richness: Effects of Proposed Tax Reforms for Germany, mimeo, Finanzwissenschaftliches Forschungsinstitut an der Universitaet zu Koeln.

Quinke, H. (2001). Erneuerung der Stichprobe des ESt-Modells des Bundesministeriums der Finanzen auf Basis der Lohn- und Einkommensteuerstatistik 1995, GMD - Forschungszentrum Informationstechnik GmbH, Technical Report.

Radner, D., Allen, R., Gonzales, M. E., Jabine, T. B. and Muller, H. J. (1980). Report on exact and statistical matching techniques, Statistical Policy Working Paper 5, Federal Committee on Statistical Methodology.

Rässler, S. (2002). *Statistical Matching*, Springer, New York [u.a.].

Rosenbaum, P. R. and Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects, *Biometrika* **70**: 41–55.

Rubin, D. (1987). *Multiple Imputation for Nonresponse in Surveys*, John Wiley & Sons, New York.

Rutherford, T. (1999). Applied General Equilibrium Modeling with MPSGE as a GAMS Subsystem: An Overview of the Modeling Framework and Syntax, *Computational Economics* **14**: 1–46.

Savard, L. (2003). Poverty and Income Distribution in a CGE-Household Micro- Simulation Model: Top-Down/Bottom Up Approach, CIRPÉE Centre interuniversitaire sur le risque, les politiques économiques et lŠemploi. Working Paper 03-43.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*, Chapman & Hall, London.

Schmidt, C. and Straubhaar, T. (1996). Bevölkerungsentwicklung und Wirtschaftswachstum - Eine Simulationsanalyse für die Schweiz, *Schweizerische Zeitschrift für Volkswirtschaft und Statistik* **132**(3): 395–414.

Sims, C. A. (1972a). Comments, *Annals of Economic and Social Measurement* **1**: 343–345.

Sims, C. A. (1972b). Rejoinder, *Annals of Economic and Social Measurement* **1**: 355–357.

Sims, C. A. (1974). Comment, *Annals of Economic and Social Measurement* **3**: 395–397.

SOEP Group (2001). The German Socio-Economic Panel (GSOEP) after more than 15 years - Overview, *Vierteljahrshefte zur Wirtschaftsforschung* **70**: 7–14.

Statistisches Bundesamt (2005). Volkswirtschaftliche Gesamtrechnungen: Input-Output-Rechnung, Fachserie 18, Reihe 2.

Sutherland, H., Taylor, R. and Gomulka, J. (2002). Combining Household Income and Expenditure Data in Policy Simulations, *Review of Income and Wealth* **48**(4): 517–536.

Train, K. (2003). *Discrete Choice Models Using Simulation*, Cambridge University Press, Cambridge.

Van Soest, A. (1995). Structural Models of Family Labor Supply: A Discrete Choice Approach, *Journal of Human Resources* **30**: 63–88.

Van Soest, A. and Das, M. (2001). Family Labor Supply and Proposed Tax Reforms in the Netherlands, *De Economist* **149(2)**: 191–218.

Wagenhals, G. (2004). Tax-benefit microsimulation models for Germany: A Survey, *IAW-Report / Institut fuer Angewandte Wirtschaftsforschung (Tübingen)* **32**(1): 55–74.

Welsch, H. (2001). Armington Elasticities and Product Diversity in the European Community: A Comparative Assessment of Four Countries, Working Paper, University of Oldenburg.

FiFo-CPE Discussion Papers

Finanzwissenschaftliche Diskussionsbeiträge

Eine Schriftenreihe des Finanzwissenschaftlichen Forschungsinstituts an der Universität zu Köln; ISSN 0945-490X.

Die Beiträge ab 1998 stehen auch als kostenloser Download zur Verfügung unter: http:// fifo-koeln.de oder http://cpe.uni-koeln.de.

Discussions Papers from 1998 onwards can be downloaded free of charge from: http:// fifo-koeln.de or http://cpe.uni-koeln.de.

93-1 Ewringmann, D.: Ökologische Steuerreform? Vergriffen.

93-2 Gawel, E.: Bundesergänzungszuweisungen als Instrument eines rationalen Finanzausgleichs. Vergriffen.

93-3 Ewringmann, D./Gawel, E./Hansmeyer, K.-H.: Die Abwasserabgabe vor der vierten Novelle: Abschied vom gewässergütepolitischen Lenkungs- und Anreizinstrument, 2. Aufl. Vergriffen.

93-4 Gawel, E.: Neuere Entwicklungen der Umweltökonomik. Vergriffen.

93-5 Gawel, E.: Marktliche und außermarktliche Allokation in staatlich regulierten Umweltmedien: Das Problem der Primärallokation durch Recht. Vergriffen.

94-1 Gawel, E.: Staatliche Steuerung durch Umweltverwaltungsrecht — eine ökonomische Analyse. Vergriffen.

94-2 Gawel, E.: Zur Neuen Politischen Ökonomie der Umweltabgabe. Vergriffen.

94-3 Bizer, K./Scholl, R.: Der Beitrag der Indirekteinleiterabgabe zur Reinhaltung von Klärschlamm. Vergriffen.

94-4 Bizer, K.: Flächenbesteuerung mit ökologischen Lenkungswirkungen. Vergriffen.

95-1 Scholl, R.: Verhaltensanreize der Abwasserabgabe: eine Untersuchung der Tarifstruktur der Abwasserabgabe. 6,50 EUR.

95-2 Kitterer, W: Intergenerative Belastungsrechnungen („Generational Accounting“) - Ein Maßstab für die Belastung zukünftiger Generationen? 7,50 EUR.

96-1 Ewringmann, D./Linscheidt, B./Truger, A.: Nationale Energiebesteuerung: Ausgestaltung und Aufkommensverwendung. 10,00 EUR.

96-2 Ewringmann, D./Scholl. R.: Zur fünften Novellierung der Abwasserabgabe; Meßlösung und sonst nichts? 7,50 EUR.

97-1 Braun, St./Kambeck, R.: Reform der Einkommensteuer. Neugestaltung des Steuertarifs. 7,50 EUR.

97-2 Linscheidt, B./Linnemann, L.: Wirkungen einer ökologischen Steuerreform – eine vergleichende Analyse der Modellsimulationen von DIW und RWI. 5,00 EUR.

97-3 Bizer, K./Joeris, D.: Bodenrichtwerte als Bemessungsgrundlage für eine reformierte Grundsteuer. 7,50 EUR.

98-1 Kitterer, W.: Langfristige Wirkungen öffentlicher Investitionen - theoretische und empirische Aspekte. 6,00 EUR.

98-2 Rhee, P.-W.: Fiskale Illusion und Glory Seeking am Beispiel Koreas (1960-1987). 5,00 EUR.

98-3 Bizer, K.: A land use tax: greening the property tax system. 5,00 EUR.

00-1 Thöne, M.: Ein Selbstbehalt im Länderfinanzausgleich?. 6,00 EUR.

00-2 Braun, S., Kitterer, W.: Umwelt-, Beschäftigungs- und Wohlfahrtswirkungen einer ökologischen Steuerreform: eine dynamische Simulationsanalyse unter besonderer Berücksichtigung der Anpassungsprozesse im Übergang. 7,50 EUR.

02-1 Kitterer, W.: Die Ausgestaltung der Mittelzuweisungen im Solidarpakt II. 5,00 EUR.

05-1 Peichl, A.: Die Evaluation von Steuerreformen durch Simulationsmodelle. 8,00 EUR.

05-2 Heilmann, S.: Abgaben- und Mengenlösungen im Klimaschutz: die Interaktion von europäischem Emissionshandel und deutscher Ökosteuer. 8,00 EUR.

05-3 Fuest, C., Peichl, A., Schaefer, T.: Dokumentation FiFoSiM: Integriertes Steuer-Transfer-Mikrosimulations- und CGE-Modell. 8,00 EUR.

06-1 Fuest, C., Peichl, A., Schaefer, T.: Führt Steuervereinfachung zu einer „gerechteren“ Einkommensverteilung? Eine empirische Analyse für Deutschland. 6,00 EUR.

06-2 Bergs, C., Peichl, A.: Numerische Gleichgewichtsmodelle - Grundlagen und Anwendungsgebiete. 6,00 EUR.

06-3 Thöne, M.: Eine neue Grundsteuer – Nur Anhängsel der Gemeindesteuerreform? 6,00 EUR.

06-4 Mackscheidt, K.: Über die Leistungskurve und die Besoldungsentwicklung im Laufe des Lebens. 6,00 EUR.

06-5 Fuest, C., Peichl, A., Schaefer, T.: Does tax simplification yield more equity and efficiency? An empirical analysis for Germany. 6,00 EUR.

06-6 Fuest, C., Peichl, A., Schaefer, T.: Die Flat Tax: Wer gewinnt? Wer verliert? Eine empirische Analyse für Deutschland. 6,00 EUR.

06-7 Kitterer, W., Finken, J.: Zur Nachhaltigkeit der Länderhaushalte – eine empirische Analyse. 6,00 EUR.

06-8 Bergs, C., Fuest, C., Peichl, A., Schaefer, T.: Reformoptionen der Familienbesteuerung: Aufkommens-, Verteilungs- und Arbeitsangebotseffekte. 6,00 EUR.

06-9 Ochmann, R., Peichl, A.: Measuring distributional effects of fiscal reforms. 10,00 EUR.

06-10 Peichl, A., Schaefer, T.: Documentation FiFoSiM: Integrated tax benefit microsimulation and CGE model. 8,00 EUR.

06-11 Peichl, A., Schaefer, T., Scheicher, C.: Measuring Richness and Poverty. A micro data application to Germany and the EU-15. 6,00 EUR.